



Statistics/Data Analysis

MP - Parallel Edition

(R)

14.0

Copyright 1985-2015 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATATA-PC <http://www.stata.com>
979-696-4600 stata@stata.com
979-696-4601 (fax)

Single-user 8-core Stata perpetual license:

Serial number: 10699393

Licensed to: Bobby W. Chung

University of Illinois at Urbana-Champaign

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . doedit "C:\Users\Bobby W. Chung\OneDrive\Project E\Project\Occupational License\Archive\Table4_Pane
2 . do "C:\Users\Bobby W. Chung\OneDrive\Project E\Project\Occupational License\Archive\Table4_Pane
3 . ***User should have five .dta files ready including:
4 . *Core file from Wave 13-16 (wave13-16_FULL.dta)
5 . *License file from the Topcial module in Wave 13 (Module.dta)
6 . *Data defining ban and non-ban licensing (mandatory_unique.dta)
7 . *incarceration disparity in race in 2014 (disparity2014.dta)
8 . *skill indexes from O*NET (Skills(importance).dta)
9 .
10 . ***Line 10 to line 327 are all about data cleaning
11 . ***User should start with the core file and run the folloing code
12 . set matsize 1500
13 . clear
14 .
15 . use "C:\Users\Bobby W. Chung\Desktop\SIPPdata\wave13-16_FULL.dta"
16 . keep if swave==13
    (791,221 observations deleted)
17 .
18 . *****Merge with license
19 . merge m:1 lgtkey using "C:\Users\Bobby W. Chung\Desktop\SIPPdata\Module.dta", generate(_license
```

Result	# of obs.	
not matched	1,814	
from master	1,814	(_licensemerge==1)
from using	0	(_licensemerge==2)
matched	302,379	(_licensemerge==3)

```

20 .
21 .
22 . merge m:1 lgtkey using "C:\Users\Bobby W. Chung\Desktop\SIPPdata\Wave2_topical", generate(edumerge)

```

Result	# of obs.	
not matched	83,448	
from master	48,966	(edumerge==1)
from using	34,482	(edumerge==2)
matched	255,227	(edumerge==3)

```

23 . drop if edumerge==2
    (34,482 observations deleted)

24 . gen math=0

25 . replace math=1 if ecourse1==1
    (102,291 real changes made)

26 . gen sci=0

27 . replace sci=1 if ecourse2==1
    (99,592 real changes made)

28 . gen eng=0

29 . replace eng=1 if ecourse3==1
    (148,508 real changes made)

30 . *****Data Management*****
31 . drop if tage<18
    (71,482 observations deleted)

32 . drop if tage >64
    (49,165 observations deleted)

33 .
34 . * Work hour<1
35 . drop if ejbhrs1<1
    (77,858 observations deleted)

36 .
37 .
38 . *delege useless license data
39 . drop if iprocert <1
    (9,461 observations deleted)

40 .
41 . *Delete imptued license

```

```
42 . drop if iwhopcer <0
    (366 observations deleted)

43 .
44 . * Define demographic
45 . generate female =1

46 . replace female =0 if esex==1
    (47,801 real changes made)

47 . drop if female==1
    (48,060 observations deleted)

48 .
49 . generate black=0

50 . replace black=1 if erace==2
    (4,325 real changes made)

51 .
52 . generate hispanic=0

53 . replace hispanic=1 if eorigin==1 & black==0
    (6,187 real changes made)

54 .
55 . *Create hourly wage
56 . generate long wage = (tpmsum1)/4/ejbhrs1

57 . drop if wage>100
    (685 observations deleted)

58 . drop if wage<5
    (4,047 observations deleted)

59 . gen lwage =log(wage)

60 . generate union=1

61 . replace union=0 if eunion1==2
    (37,332 real changes made)

62 .
63 . *Delete imputed union data
64 . drop if eunion1 <0
    (4 observations deleted)

65 .
66 . *Drop race other than white and black
67 . gen other=0
```

```
68 . replace other=1 if erace>2
    (3,246 real changes made)

69 .
70 . ***Merging with Module data of license (skip)
71 . ***Create license dummy (skip)
72 .
73 . *Delete imputed licensure source
74 . drop if iwhopcer <0
    (0 observations deleted)

75 .
76 .
77 . *Create gov't summy
78 . generate govt=0

79 . replace govt=1 if inrange(eclwrk1, 3,5)
    (6,683 real changes made)

80 .
81 . *service worker?
82 . gen service=0

83 . replace service=1 if inrange(ejbind1,6870,9290)
    (17,093 real changes made)

84 .
85 . *Create education dummy
86 . generate hs =0

87 . replace hs=1 if eeducate==39
    (11,397 real changes made)

88 . generate somecollege =0

89 . replace somecollege=1 if inrange(eeducate,40,43)
    (14,748 real changes made)

90 . generate college=0

91 . replace college=1 if eeducate ==44
    (9,228 real changes made)

92 . generate postgrad=0

93 . replace postgrad=1 if eeducate >44
    (4,531 real changes made)

94 .
95 . *Create self-employment and big company dummy
96 . generate self_emp=1
```

```

97 . replace self_emp=0 if ebuscntr<0
    (41,861 real changes made)

98 .
99 . gen age2=tage*tage

100 . rename tage age

101 .
102 .
103 . *Ensure that the license variable does not contain certified
104 . replace license=0 if cert==1
    (3,847 real changes made)

105 .
106 . *Ensure the license indicator include only license being required by job
107 . gen licensenorequire=0

108 . replace licensenorequire=1 if irjpcert!=1 & license==1
    (1,298 real changes made)

109 . replace license=0 if irjpcert>1
    (1,637 real changes made)

110 .
111 .
112 . ***** Redefine licensing by 50/50 rule
113 . egen licensecheck=mean(license), by(tfipsst tjboccl)

114 . replace license=0 if licensecheck<0.5
    (1,845 real changes made)

115 . replace license=1 if licensecheck>=0.5
    (1,245 real changes made)

116 .
117 .
118 .             *Merge with ban data
119 .             gen state_occ=string(tfipsst)+"-"+string(tjboccl) if license==1
    (38,163 missing values generated)

120 .
121 .             merge m:1 state_occ using "C:\Users\Bobby W. Chung\OneDrive\Project E\Project\Occupatio
> ate(_felonymerge)

```

Result	# of obs.	
not matched	43,249	
from master	41,014	(_felonymerge==1)
from using	2,235	(_felonymerge==2)
matched	2,051	(_felonymerge==3)

```

122 .
123 .     drop if _felonymerge==2
      (2,235 observations deleted)

124 .     replace permбан=0 if _felonymerge==1
      (41,014 real changes made)

125 .     replace felony=0 if _felonymerge==1
      (41,014 real changes made)

126 .
127 .     rename permбан Ban

128 .
129 .     ***"permбан" only contains permanent ban
130 .     ***"felony" contains both temporary and permanent ban
131 .
132 .     keep if apmsum1==0 & inrange(wage,5,100)
      (4,893 observations deleted)

133 .
134 .
135 .     gen person=0

136 .     replace person=1 if iwhypcer==2
      (397 real changes made)

137 .
138 .     ***** Race selection
139 .     drop if other==1
      (2,928 observations deleted)

140 .     drop if hispanic==1 & black==0
      (4,908 observations deleted)

141 .
142 .
143 .     *****
144 .     **** Main Table
145 .     *****
146 .     replace quality=100 if missing(quality)
      (28,786 real changes made)

147 .
148 .     local controls age age2 hs somecollege college postgrad union govt self_em service

149 .     *quality
150 .     *eng math sci person
151 .
152 .     **Add [pweight=wpfinwgt] for weighted OLS
153 .

```

```

154 .
155 .     gen boxyes=0

156 .     replace boxyes=1 if inlist(tfipsst,6,8,9,11,15,25,27,35)
      (4,716 real changes made)

157 .
158 .     gen BTBban=Ban*boxyes

159 .     replace Ban=0 if BTBban==1
      (246 real changes made)

160 .     gen nonBTBban=Ban

161 .
162 .
163 .     gen con_edu=0

164 . replace con_edu=1 if icdpcert==1
      (5,178 real changes made)

165 .
166 .
167 .     **** All license
168 .
169 . quietly areg lwage black license cert licensenorequire `controls' i.tfipsst i.rhcalmn if apn
      > st) absorb(tjboccl)

170 .     psacalc delta license, beta(0) rmax(0.61)

```

—— Bound Estimate ——		
delta	1.98528	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.17752	0.010
Controlled	0.04483	0.560
—— Other Inputs ——		
R_max	0.610	
Beta	0.000000	
Unr. Controls		

```

171 .     psacalc delta license, beta(0) rmax(0.72)

```

—— Bound Estimate ——		
delta	0.62282	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.17752	0.010
Controlled	0.04483	0.560

— Other Inputs —	
R_max	0.720
Beta	0.000000
Unr. Controls	

172 . psacalc delta license, beta(0) rmax(1)

— Bound Estimate —	
delta	0.22674

— Inputs from Regressions —		
	Coeff.	R-Squared
Uncontrolled	0.17752	0.010
Controlled	0.04483	0.560

— Other Inputs —	
R_max	1.000
Beta	0.000000
Unr. Controls	

173 .

174 .

175 . ***** BY race

176 . quietly areg lwage license cert licensenorequire `controls' i.tfipsst i.rhcalmn if apmsum1==0
> fipsst) absorb(tjboc1)

177 . psacalc delta license, beta(0) rmax(0.61)

— Bound Estimate —	
delta	1.92048

— Inputs from Regressions —		
	Coeff.	R-Squared
Uncontrolled	0.15110	0.008
Controlled	0.03941	0.558

— Other Inputs —	
R_max	0.610
Beta	0.000000
Unr. Controls	

178 . psacalc delta license, beta(0) rmax(0.72)

—— Bound Estimate ——		
delta	0.62244	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.15110	0.008
Controlled	0.03941	0.558
—— Other Inputs ——		
R_max	0.720	
Beta	0.000000	
Unr. Controls		

179 . psacalc delta license, beta(0) rmax(1)

—— Bound Estimate ——		
delta	0.22880	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.15110	0.008
Controlled	0.03941	0.558
—— Other Inputs ——		
R_max	1.000	
Beta	0.000000	
Unr. Controls		

180 .

181 . quietly areg lwage license cert licensenorequire `controls' i.tfipsst i.rhcalmn if apmsum1==0
> fipsst) absorb(tjboc1)

182 . psacalc delta license, beta(0) rmax(0.75)

—— Bound Estimate ——		
delta	3.05947	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.28021	0.015
Controlled	0.15894	0.685

— Other Inputs —	
R_max	0.750
Beta	0.000000
Unr. Controls	

183 . psacalc delta license, beta(0) rmax(0.89)

— Bound Estimate —	
delta	0.98627

— Inputs from Regressions —		
	Coeff.	R-Squared
Uncontrolled	0.28021	0.015
Controlled	0.15894	0.685

— Other Inputs —	
R_max	0.890
Beta	0.000000
Unr. Controls	

184 . psacalc delta license, beta(0) rmax(1)

— Bound Estimate —	
delta	0.64360

— Inputs from Regressions —		
	Coeff.	R-Squared
Uncontrolled	0.28021	0.015
Controlled	0.15894	0.685

— Other Inputs —	
R_max	1.000
Beta	0.000000
Unr. Controls	

185 .

186 .

```

187 .
    end of do-file

188 . do "C:\Users\Bobby W. Chung\OneDrive\Project E\Project\Occupational License\Archive\Table4_Pane

189 . ***User should have three .dta files ready including:
190 . *Core file from Wave 13-16 (wave13-16_FULLL.dta)
191 . *License file from the Topcial module in Wave 13 (Module.dta)
192 . *Data defining ban and non-ban licensing (mandatory_unique.dta)
193 . *Wave 2 that contains history of advanced courses taken (Wave2_topical.dta)
194 .
195 .
196 . ***Line 10 to line 327 are all about data cleaning
197 . ***User should start with the core file and run the folloing code
198 . set matsize 1500

```

```

199 . clear

200 .
201 . use "C:\Users\Bobby W. Chung\Desktop\SIPPdata\wave13-16_FULLL.dta"

```

```

202 . keep if swave==13
    (791,221 observations deleted)

```

```

203 .
204 . *****Merge with license
205 . merge m:1 lgtkey using "C:\Users\Bobby W. Chung\Desktop\SIPPdata\Module.dta", generate(_license

```

Result	# of obs.	
not matched	1,814	
from master	1,814	(_licensemerge==1)
from using	0	(_licensemerge==2)
matched	302,379	(_licensemerge==3)

```

206 .
207 .
208 . merge m:1 lgtkey using "C:\Users\Bobby W. Chung\Desktop\SIPPdata\Wave2_topical", generate(edume

```

Result	# of obs.	
not matched	83,448	
from master	48,966	(edumerge==1)
from using	34,482	(edumerge==2)
matched	255,227	(edumerge==3)

```

209 . drop if edumerge==2
    (34,482 observations deleted)

```

```
210 . gen math=0

211 . replace math=1 if ecourse1==1
    (102,291 real changes made)

212 . gen sci=0

213 . replace sci=1 if ecourse2==1
    (99,592 real changes made)

214 . gen eng=0

215 . replace eng=1 if ecourse3==1
    (148,508 real changes made)

216 . *****Data Management*****
217 . drop if tage<18
    (71,482 observations deleted)

218 . drop if tage >64
    (49,165 observations deleted)

219 .
220 . * Work hour<1
221 . drop if ejbhrs1<1
    (77,858 observations deleted)

222 .
223 .
224 . *delege useless license data
225 . drop if iprocert <1
    (9,461 observations deleted)

226 .
227 . *Delete imptued license
228 . drop if iwhopcer <0
    (366 observations deleted)

229 .
230 . * Define demographic
231 . generate female =1

232 . replace female =0 if esex==1
    (47,801 real changes made)

233 . drop if female==1
    (48,060 observations deleted)

234 .
235 . generate black=0
```

```
236 . replace black=1 if erace==2
      (4,325 real changes made)

237 .
238 . generate hispanic=0

239 . replace hispanic=1 if eorigin==1 & black==0
      (6,187 real changes made)

240 .
241 . *Create hourly wage
242 . generate long wage = (tpmsum1)/4/ejbhrs1

243 . drop if wage>100
      (685 observations deleted)

244 . drop if wage<5
      (4,047 observations deleted)

245 . gen lwage =log(wage)

246 . generate union=1

247 . replace union=0 if eunion1==2
      (37,332 real changes made)

248 .
249 . *Delete imputed union data
250 . drop if eunion1 <0
      (4 observations deleted)

251 .
252 . *Drop race other than white and black
253 . gen other=0

254 . replace other=1 if erace>2
      (3,246 real changes made)

255 .
256 . ***Merging with Module data of license (skip)
257 . ***Create license dummy (skip)
258 .
259 . *Delete imputed licensure source
260 . drop if iwhopcer <0
      (0 observations deleted)

261 .
262 .
263 . *Create gov't summy
264 . generate govt=0
```

```
265 . replace govt=1 if inrange(eclwrk1, 3,5)
    (6,683 real changes made)

266 .
267 . *service worker?
268 . gen service=0

269 . replace service=1 if inrange(ejbind1,6870,9290)
    (17,093 real changes made)

270 .
271 . *Create education dummy
272 . generate hs =0

273 . replace hs=1 if eeducate==39
    (11,397 real changes made)

274 . generate somecollege =0

275 . replace somecollege=1 if inrange(eeducate,40,43)
    (14,748 real changes made)

276 . generate college=0

277 . replace college=1 if eeducate ==44
    (9,228 real changes made)

278 . generate postgrad=0

279 . replace postgrad=1 if eeducate >44
    (4,531 real changes made)

280 .
281 . *Create self-employment and big company dummy
282 . generate self_emp=1

283 . replace self_emp=0 if ebuscntr<0
    (41,861 real changes made)

284 .
285 . gen age2=tage*tage

286 . rename tage age

287 .
288 .
289 . *Ensure that the license variable does not contain certified
290 . replace license=0 if cert==1
    (3,847 real changes made)

291 .
```

```

292 . *Ensure the license indicator include only license being required by job
293 . gen licensenorequire=0

294 . replace licensenorequire=1 if irjpcert!=1 & license==1
    (1,298 real changes made)

295 . replace license=0 if irjpcert>1
    (1,637 real changes made)

296 .
297 .
298 . ***** Redefine licensing by 50/50 rule
299 . egen licensecheck=mean(license), by(tfipsst tjboccl)

300 . replace license=0 if licensecheck<0.5
    (1,845 real changes made)

301 . replace license=1 if licensecheck>=0.5
    (1,245 real changes made)

302 .
303 .
304 .             *Merge with ban data
305 .             gen state_occ=string(tfipsst)+"-"+string(tjboccl) if license==1
    (38,163 missing values generated)

306 .
307 .             merge m:1 state_occ using "C:\Users\Bobby W. Chung\OneDrive\Project E\Project\Occupatio
> ate(_felonymerge)

      Result                # of obs.
-----
not matched                43,249
  from master              41,014  (_felonymerge==1)
  from using               2,235  (_felonymerge==2)

matched                    2,051  (_felonymerge==3)
-----

308 .
309 .             drop if _felonymerge==2
    (2,235 observations deleted)

310 .             replace permбан=0 if _felonymerge==1
    (41,014 real changes made)

311 .             replace felony=0 if _felonymerge==1
    (41,014 real changes made)

312 .
313 .             rename permбан Ban

```

```

314 .
315 .     ***"permban" only contains permanent ban
316 .     ***"felony" contains both temporary and permanent ban
317 .
318 .     keep if apmsum1==0 & inrange(wage,5,100)
      (4,893 observations deleted)

319 .
320 .
321 .     gen person=0

322 .     replace person=1 if iwhypcer==2
      (397 real changes made)

323 .
324 .     ***** Race selection
325 .     drop if other==1
      (2,928 observations deleted)

326 .     drop if hispanic==1 & black==0
      (4,908 observations deleted)

327 .
328 .
329 .     *****
330 .     **** Main Table
331 .     *****
332 .     replace quality=100 if missing(quality)
      (28,786 real changes made)

333 .
334 .     local controls  age age2  hs somecollege college postgrad union govt self_em service

335 .     *quality
336 .     *eng math sci person
337 .
338 .     **Add [pweight=wpfinwgt] for weighted OLS
339 .
340 .
341 .     gen boxyes=0

342 .     replace boxyes=1 if inlist(tfipsst,6,8,9,11,15,25,27,35)
      (4,716 real changes made)

343 .
344 .     gen BTBban=Ban*boxyes

345 .     replace Ban=0 if BTBban==1
      (246 real changes made)

346 .     gen nonBTBban=Ban

```

```

347 .
348 .
349 .         gen con_edu=0

350 . replace con_edu=1 if icdpcert==1
    (5,178 real changes made)

351 .
352 .
353 .         **** All license
354 .
355 . quietly areg lwage black license cert licensenorequire `controls' i.tfipsst i.rhcalmn if apn
    > st) absorb(tjboccl)

356 .         psacalc delta license, beta(0) rmax(0.61)

```

—— Bound Estimate ——		
delta	1.98528	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.17752	0.010
Controlled	0.04483	0.560
—— Other Inputs ——		
R_max	0.610	
Beta	0.000000	
Unr. Controls		

```

357 .         psacalc delta license, beta(0) rmax(0.72)

```

—— Bound Estimate ——		
delta	0.62282	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.17752	0.010
Controlled	0.04483	0.560
—— Other Inputs ——		
R_max	0.720	
Beta	0.000000	
Unr. Controls		

358 . psacalc delta license, beta(0) rmax(1)

—— Bound Estimate ——		
delta	0.22674	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.17752	0.010
Controlled	0.04483	0.560
—— Other Inputs ——		
R_max	1.000	
Beta	0.000000	
Unr. Controls		

359 .

360 .

361 . ***** BY race

362 . quietly areg lwage license cert licensenorequire `controls' i.tfipsst i.rhcalmn if apmsum1==0
> fipsst) absorb(tjboc1)

363 . psacalc delta license, beta(0) rmax(0.61)

—— Bound Estimate ——		
delta	1.92048	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.15110	0.008
Controlled	0.03941	0.558
—— Other Inputs ——		
R_max	0.610	
Beta	0.000000	
Unr. Controls		

364 . psacalc delta license, beta(0) rmax(0.72)

—— Bound Estimate ——		
delta	0.62244	
—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.15110	0.008
Controlled	0.03941	0.558

— Other Inputs —	
R_max	0.720
Beta	0.000000
Unr. Controls	

365 . psacalc delta license, beta(0) rmax(1)

— Bound Estimate —	
delta	0.22880

— Inputs from Regressions —		
	Coeff.	R-Squared
Uncontrolled	0.15110	0.008
Controlled	0.03941	0.558

— Other Inputs —	
R_max	1.000
Beta	0.000000
Unr. Controls	

366 .
 367 . quietly areg lwage license cert licensenorequire `controls' i.tfipsst i.rhcalmn if apmsum1==0
 > fipsst) absorb(tjboccl)

368 . psacalc delta license, beta(0) rmax(0.75)

— Bound Estimate —	
delta	3.05947

— Inputs from Regressions —		
	Coeff.	R-Squared
Uncontrolled	0.28021	0.015
Controlled	0.15894	0.685

— Other Inputs —	
R_max	0.750
Beta	0.000000
Unr. Controls	

369 . psacalc delta license, beta(0) rmax(0.89)

—— Bound Estimate ——		
delta	0.98627	

—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.28021	0.015
Controlled	0.15894	0.685

—— Other Inputs ——	
R_max	0.890
Beta	0.000000
Unr. Controls	

370 . psacalc delta license, beta(0) rmax(1)

—— Bound Estimate ——		
delta	0.64360	

—— Inputs from Regressions ——		
	Coeff.	R-Squared
Uncontrolled	0.28021	0.015
Controlled	0.15894	0.685

—— Other Inputs ——	
R_max	1.000
Beta	0.000000
Unr. Controls	

371 .

372 .

373 .

end of do-file

374 .